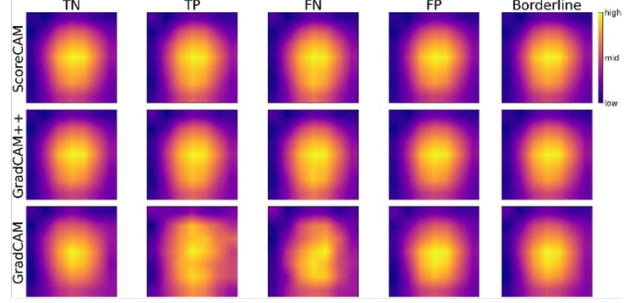
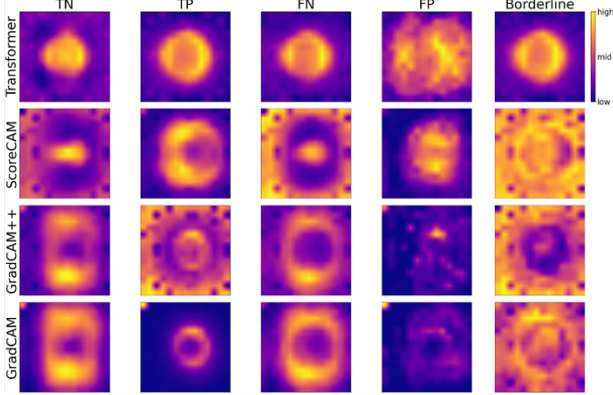


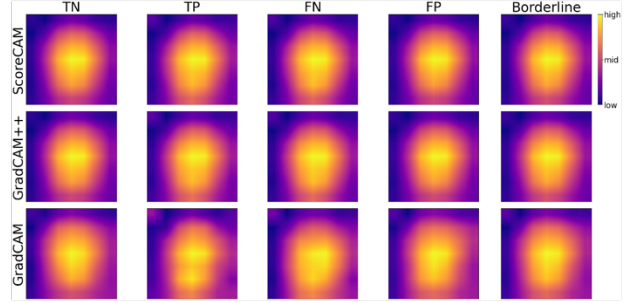
(a) DeiT attention maps (Model 4)



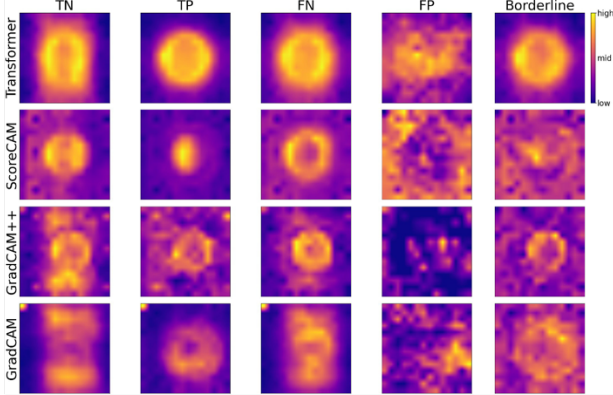
(b) ResNet-50 saliency maps (Model 4)



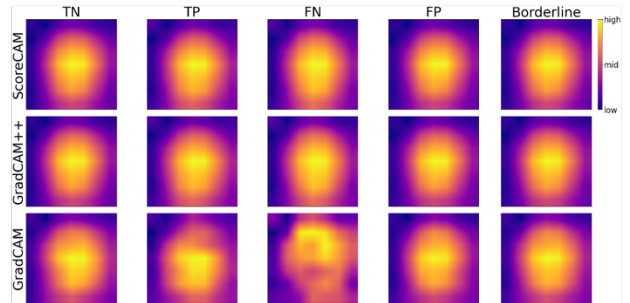
(c) DeiT attention maps (Model 2)



(d) ResNet-50 saliency maps (Model 2)



(e) DeiT attention maps (Model 5)



(f) ResNet-50 saliency maps (Model 5)

Supplemental Figure 4. The averages of attention/saliency maps are compared between DeiT and ResNet-50 for Model 4, Model 2, and Model 5 in the LAG dataset. We categorize the results based on model outcomes (TP, TN, FP, FN, borderline). The borderline cases are those where the model confidence in its answer is between 0.3 and 0.7.